International Journal of Sciences: Basic and Applied Research (IJSBAR)

ISSN 2307-4531
(Print & Online)

http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

Promulgation of Mathematical Test Characteristics-Based to be a Benchmark on Mathematical Learning Outcome

Dr. Waminton Rajagukguk*, M.Pd.

Associate Professor, Medan State University

Email: warajagukguk@gmail.com

Abstract

This research is intended to promulgate a form of mathematical test characteristics-based in order to be a benchmark on mathematical learning outcomes. The research took samples randomly out of 300 students by "random sampling design." The test constructed comprising multiple choice with 40 items of questions. The research data analysis by validity, reliability, difficult index, discriminative index test. After looking into the implementation of test quality in fact: (a) 33 items of questions were valid. (b) coeficient reliability of 0.86 is a higher degree reliability. (c) difficult index 20% of tough category, 65% of fair tough category, 15% of easy category (d) by this 33 items of question within good discrimative index enables to discern students who are yet mastery learning and who are not at all. Whereby this three consecutive trials experiment inferred that 33 items of questions that existed from the third step test is properly shifted to become a standard or benchmark to the mathematical characterics-based test.

Keywords: Benchmark Test; Evaluation; Mathematical Characteristics-Based; Validity; Reliability; Difficult Index; Discriminative Index.

1. Introdction

The necessity of evaluation in order to govern educational quality, by means of planning and implementing entailed to get the learning outcome which is well-suited with stipulated educational purpose [11].

---

* Corresponding author.

E-mail address: warajagukguk@gmail.com.

Regarding with teachers this very learning outcome is not remained listed only to represent an accountable report upom their superior or just remained as a handy grade slip, but foremost its considered as benchmark to be self introspective in manner of how well had the learning performance being done. For some students will take it as a stepping stone or sekf-assessment to make all future exerts better. In some other students will may be bear the path conciously about limited cognitive ability. To the parents, their kid's learning outcome is being a reference of good, fair, poor, and will do something to work out for it. Considering educational management, its noteworthy becoming an evaluative instrument focus to further improvement which applicable to both students and teachers.

Evaluation over learning outcome can be optimally used if it is implemented by principled assessment as a wholistic, continuity, goal-oriented, objective assessed, and overt judgement of beneficial aspects [12]. Wholistic evaluation means the collected informations include a whole personal aspect, knowledge, attitude, prowess. Objective evaluation is conducted by complying with the set forth policies on evaluational criteria.

Thus evaluation over mathematical learning outcome, a test is a requisite component to discern precisely mathematical learning teaching process quality. Test score is a sequel that is proceeded by a test or examination or quizz is explicitly to depict students' intelligential achievement. Then so-called intelligence means that how students enabled to comprehend mathematical objects like mathematical facts, prowess, consepts, principles [8]. A student being considered good at facts, if that student is coherently able to describe well about the facts themselves and to imply implementation into any situation. Good at mathematical concepts referred to if a student can exchange intangible ideas into tangible things in a daily lives. Meanwhile a student that is good at mathematical principles is how student becoming good at mathematical axiom, postulates, theorem. The reasons why the assessment of mathematical learning outcome entailed to ponder about mathematical characteristic so that the test conducted might obviously to figure out mathematical well-done benchmark [13]. Now the question is how to devise well the form of a benchmark test that is able to determine mathematical learning outcome.

There is three applicable terms of evaluation, that is: test, measurement, assessment. Test is a factor to estimate students' intelligential magnitude indirectly, which is by interactive response on stimulus or question [5]. A test is also defineable to collect informative characteristic of an object. This object can imply of students' intellectual ability, attitude, interest, and motivation. A test as considered as feedback on participants' responses on some questions can portray an intelligential adeptness in a certain thing. Measurement is definable as a figure of an establishment process or characteristic of a certain direction [6]. Evaluation has broader comprehension than valuation, due an evaluation is a process or activity to select, collect, analyze, and present information which basicly used to make decision to arrange next program and delineate it [10].

A test on mathematical learning outcome which is based on learnt objects in mathematics referred facts, skills, concepts, principles, beyond learning mathematics there are transfer of learning, inquiry ability, problem solving, self-discipline, appreciative mathematical structure [7].

Mathematical facts is any conventions (deals) in mathematics. Facts are learnt by any techniques without

thinking (*rote learning*) like memorizing, exercising, practicing, periodical test, games, contest [4]. Mathematical prowess is an operation and procedure, where a student is expected being able to accomplish quickly and correctly. Many skills are clearly explainable by a bunch of regulations and instructions or by a series of sequencial procedures which so-called algorithm. Amongst mathematical prowess which expected must be good at by all people is: long consecutive devision, fractional addition and decimal fractional multiplication. Skill is learned by demonstrative and any exercises and practicum like worksheet, work on whiteboard, study group, etc. The students are deemed good at skill, when they are aptly to demonstrate skills precisely and correctly in accomplishing all kind of questions, or excercising that skills in any situations [2].

Mathematical concepts is an intangible idea which facilitates people easily to classify objects or events, and to identify whether the objects or events are being an instance or not, equation, inequality, triangle, cube, radius, and exponent, etc. Concepts are learnable by defenitions or by direct observing, whereas students learn how to classify objects area to be triangle sets, circumference, squares, etc., but just a few of kids who are smart to describe about triangle concepts. A concep can be learned by listening, viewing, touching, discussing, or think about any instances and not from the concept itself, then arguing about instance or not instance [1].

Mathematical principle is a most complex mathematical object. Principle is a series of concepts with which related among the concepts themselves. Statement: " two triangles are congruous if two sides and the wedge angles are superimposed," and "square of hypotenuse of right triangle equals to square of both side elbows". Its a triangled principle and Phytagoras principle. Principles can be unveiled by scientific inquiry process, guided invention, discussion group, applying problem solving strategy, and demonstrative. The students had learnt principles if the students have coherently incisively to determine concepts within principles, to place these concepts into right relation between one and each other and bring the principle out of a certain condition. Whenever mathematical learning purpose is high comprehensive about concepts, the students should have been unconcealed about learning concepts with deeply enough yet they are enabled to define to their colleagues students and immediately to adapt in a certain situation.

In order to meet this mathematical learning purpose entailed some good competencies that students supposedly to qualify for. These competencies are again broken down into competent set standard. This set standard is detailed in basic competence, indicator, and main material, of each aspects. Referred to standard and basic competency in mathematical sphere that should be regained by students in the scope of mathematical material is algebra, geometry, trigonometry, and calculus [3]. In this case mathematical characteristic competency detail falls into:

1. Having intangible objective study about relation of facts, operation, concepts, and priciples. Facts are mathematical agreement and convention which are used to be disclosed by certain symbols. Operation is a processing count of algebra and mathematics. Relation is connection between two or more elements. Concepts are intangible ideas which are used to classify or categorize a bunch of objects whether certain objects are as instances concepts or not. Principle is the complex mathematical objects in a form of axioms, principles (postulates), characteristics, etc.
2. Focusing on some deals whereas symbols and mathematical terms denote essential agreements and

conventions.

3. Deductive thinking figure whereas original thoughts are out of universality applied or directed to spesific things.
4. Consistent or adherently to its systems, mathematics has many various systems which formed by axioms and theorems.
5. Having any unexplained symbols, that mathematics supposed to have embedded the explained symbols then so called mathematical model.
6. Mathematics is an art of creativity which entailed imaginative, intuitive, comprehensive.
7. An act of problem solving.

To construct a test pattern of mathematical learning outcome must have required good discriminative index, difficult index, and high reliability where mathematical characteristic-based employed as reference in constructing test. The valuation of mathematical learning outcome aspects that should have come into considerations are:

a. Comprehensive concepts denote competency which are exposed by students in understanding concepts and subject to flexible procedures, accurateness, efficiency, preciseness and these things delineated with:

1.1. Review a concept
1.2. Exemplify and not an instance from a concept itself
1.3. Provide any mathematical presentation
1.4. Develop requisite or adequate condition of a concept
1.5. Employ, use, and choose any procedure or certain operation
1.6. Apply the concepts

b. Reasoning and communicative mathematical thoughts elicited in:

1.1. Providing verbal, written mathematical statement, graphic, diagram
1.2. Doing mathematical manipulation
1.3. Drawing a conclusion, compiling evidence, justifying or verifying solution
1.4. Drawing a conclusion from a statement
1.5. Checking argumentative validity out
1.6. Finding pattern to establish generalization

c. Problem solving which is evoked by competence in mathematical reasoning and communicative thoughts which included:

1.1. Organize data and select informations which are relevant in problem solving
1.2. Provide mathematical problem in any case
1.3. Select accurate approach and solving problem method
1.4. Evolve solving problem strategy
1.5. Create cases

1.6. Create and interpret mathematical model from any cases

1.7. Solving unprecedented or unusual problems

2. Methodology of the research

The purpose of this research is to unfold a benchmark test that which used to be a tool of measuring mathematical learning outcome that works for students, teachers and educational institutions.

In order to elicit a benchmark test by mathematical characteristic-based learning outcome which was preconceived by 4-D's Model: *Define, Design, Develop, Disseminate* [5]. The phase of Define is to specify the test requirements. The phase of Design is to get the *prototype* (test sample) as so-called Draft-A [9]. The phase of Develop is to generate final draft which is under revision of experts based on the spot experiment and useful to see validity, reliability, discriminative index, and difficult index. This research took place at Public Junior High School Medan city, in academic year of 2013-2014 who were get involved partaking to fill out the answers on the given test. The form of test was mathematical characteristics-based comprised multiple choice, and the students get involved into the test about 300 students, employing sample of *cluster random sampling*.

a. Frist, designed 40 items of questions
b. Second, the randomized Junior High Schools had an equality in condition, social environment, geographic location, mathematical teacher quality, facility.
c. Third, randomized the involved classes that which put them equally in this research.
d. Fourth, the test was implemented by gradual steps. First step involved 100 students and then took the test quality analysis. Second step involved other 100 students, applying the revised test from the test given to the first step. And the third step involved other 100 students applied the re-revised test from the test given to the second step, and formally constitute a benchmark test for mathematical characteristics-based.

The final step test of mathematical characteristics-based composition consisted validity, reliability, difficult index, discriminative index,

1. Validity

1. $0.80 < r_{xy} \leq 1.00$ very high correlation (very high validity)

2. $0.60 < r_{xy} \leq 0.80$ high correlation (high validity)

3. $0.40 < r_{xy} \leq 0.60$ fair correlation (fair validity)

4. $0.20 < r_{xy} \leq 0.40$ low correlation (poor validity)

5. $0.00 < r_{xy} \leq 0.20$ very low correlation (very poor validity)

If validity average 0.00 – 0.40 means its insignificant

2. Reliabilty

1. $0.80 < r_{II} \leq 1.00$ very high reliability
2. $0.60 < r_{II} \leq 0.80$ high reliability
3. $0.40 < r_{II} \leq 0.60$ fair reliability
4. $0.20 < r_{II} \leq 0.40$ poor reliability
5. $0.00 < r_{II} \leq 0.20$ very poor reliability

If reliability average 0.00 – 0.40 means its insignificant

3. Difficult Index

1. $0.70 < Dif \leq 1.00$ difficult test
2. $0.30 < Dif \leq 0.70$ fair difficult test
3. $0.00 < Dif \leq 0.30$ easy test

If difficult index average 0.00 – 0.30 means its insignificant

4. Discriminative Index

1. $0.40 < Dis \leq 1.00$ good discriminative index
2. $0.30 < Dis \leq 0.40$ fair good discriminative index
3. $0.00 < Dis \leq 0.30$ poor discriminative index

If discriminative index average 0.00 – 0.30 means its insignificant

3. Research results and discussion

The implementation of the first step test which consisted 40 items of questions in which difficult index, discriminative index were set for the significance. In order to identify items of questions validity employed point-biserial coefficient, and acquired 10 items of questions were invalid their number were: 7, 10, 15, 18, 22, 24, 26, 31, 36, 40. These were caused by overriding the text of items of questions were not clear. Look into difficult index says 0.6 to 0.8 meant these items of questions were unusable to measure students' mathematical learning intelligence. By discriminative index the items of questions pointed lower between 0.15 to 0.35 otherwise the items of questions were unusable to distinguish between students who had mastered the material and had not. Then coefficient reliability of 0.68 by applying Kuder-Richardson Formula or KR-21.

Afterwards in the implementation of second step test, after revised test left over of 37 items of questions, whereas the items of questions number of: 15, 24, 36 were unusable to see validity, difficult index, discriminative index. And the leftover 37 items of questions to be re-test again to another 100 students (different with 100 students in the first step). By data analysis acquired coefficient reliability of 0.74 items of questions were valid and 34 items of questions were invalid they were: 7, 10, and 31.

Table 1: Test Result Analysis of First Step, Validity, Difficult Index, Discriminative Index

Question Number	Validity	Difficult Index	Discriminative Index	Description
1	0.74	0.64	0.45	Significant
2	0.68	0.30	0.53	Significant
3	0.50	0.70	0.43	Significant
4	0.65	0.59	0.65	Significant
5	0.64	0.40	0.59	Significant
6	0.70	0.34	0.45	Significant
7	0.58	0.30	0.32	Insignificant
8	0.87	0.55	0.69	Significant
9	0.75	0.50	0.63	Significant
10	0.58	0.32	0.28	Insignificant
11	0.80	0.40	0.58	Significant
12	0.79	0.45	0.44	Significant
13	0.84	0.42	0.51	Significant
14	0.68	0.72	0.41	Significant
15	0.48	0.26	0.28	Insignificant
16	0.66	0.33	0.47	Significant
17	0.78	0.36	0.39	Significant
18	0.56	0.28	0.32	Insignificant
19	0.66	0.52	0.67	Significant
20	0.82	0.55	0.46	Significant
21	0.72	0.49	0.64	Significant
22	0.54	0.35	0.30	Insignificant
23	0.69	0.38	0.46	Significant
24	0.45	0.27	0.28	Insignificant
25	0.73	0.43	0.41	Significant
26	0.54	0.36	0.32	Insignificant
27	0.74	0.59	0.33	Significant
28	0.78	0.54	0.60	Significant
29	0.84	0.56	0.52	Significant
30	0.79	0.49	0.61	Significant
31	0.56	0.32	0.29	Insignificant
32	0.77	0.39	0.59	Significant
33	0.83	0.58	0.49	Significant

34	0.68	0.45	0.64	Significant
35	0.62	0.53	0.52	Significant
36	0.42	0.30	0.24	Insignificant
37	0.67	0.47	0.44	Significant
38	0.77	0.38	0.33	Significant
39	0.63	0.45	0.42	Significant
40	0.56	0.36	0.28	Insignificant

Table 2: Test Result Analysis of Second Step, Validity, Difficult Indes, Discriminative Index

Question Number	**Validity**	**Difficult Index**	**Discriminative Index**	**Description**
1	0.74	0.64	0.45	Significant
2	0.64	0.32	0.40	Significant
3	0.60	0.39	0.32	Signifcant
4	0.69	0.25	0.35	Significant
5	0.77	0.55	0.54	Significant
6	0.85	0.47	0.40	Significant
7	0.59	0.30	0.32	Insignificant
8	0.62	0.37	0.49	Significant
9	0.74	0.33	0.37	Significant
10	0.58	0.31	0.32	Insignificant
11	0.81	0.49	0.62	Significant
12	0.62	0.36	0.34	Significant
13	0.83	0.58	0.56	Significant
14	0.66	0.33	0.47	Significant
16	0.78	0.39	0.49	Significant
17	0.82	0.49	0.51	Significant
18	0.81	0.38	0.62	Significant
19	0.64	0.58	0.49	Significant
20	0.69	0.51	0.44	Significant
21	0.74	0.50	0.35	Significant
22	0.76	0.51	0.44	Significant
23	0.73	0.42	0.49	Significant
25	0.79	0.36	0.56	Significant
26	0.68	0.42	0.61	Significant
27	0.82	0.55	0.67	Significant

28	0.77	0.38	0.46	Significant
29	0.78	0.41	0.54	Significant
30	0.67	0.54	0.43	Significant
31	0.57	0.32	0.30	Insignificant
32	0.60	0.35	0.38	Significant
33	0.67	0.36	0.41	Significant
34	0.78	0.39	0.53	Significant
35	0.68	0.40	0.63	Significant
37	0.79	0.30	0.45	Significant
38	0.73	0.42	0.49	Significant
39	0.64	0.32	0.40	Significant
40	0.68	0.40	0.42	Significant

On this third step inplementation there were leftover 34 proper items of questions whereby the items of questions number of: 7, 10, 31 were unusable properly. Then re-retested again to another 100 students (different with students in first and second step). Obtained leftover 33 items of questions met the standard, and only the item question number 3 was unusable again because of low validity, difficult index and discriminative index. On this third step the coefficient reliability of 0.86 and the phase of good correlation between items' score and total's score (validity). Difficult index of 0.3 to 0.5 and difficult index of 0.4 to 0.8, afterall this items of questions are properly used to be a benchmark test to Junior High School.

Table 3: Test Result Analysis of Third Step, Validity, Difficult Index, Discriminative Index

Question Number	**Validity**	**Difficult Index**	**Discriminative Index**	**Description**
1	0.74	0.64	0.45	Significant
2	0.60	0.50	0.40	Significant
4	0.67	0.50	0.39	Significant
5	0.77	0.47	0.46	Significant
6	0.63	0.39	0.51	Significant
8	0.67	0.56	0.66	Significant
9	0.70	0.40	0.42	Significant
11	0.70	0.33	0.50	Significant
12	0.81	0.42	0.62	Significant
13	0.82	0.58	0.47	Significant
14	0.73	0.38	0.56	Significant
16	0.85	0.49	0.43	Significant

17	0.80	0.42	0.58	Significant
18	0.81	0.38	0.65	Significant
19	0.85	0.43	0.45	Significant
20	0.75	0.29	0.52	Significant
21	0.68	0.43	0.61	Significant
22	0.63	0.52	0.48	Significant
23	0.83	0.48	0.42	Significant
25	0.66	0.36	0.66	Significant
26	0.77	0.44	0.59	Significant
27	0.81	0.53	0.47	Significant
28	0.79	0.37	0.56	Significant
29	0.65	0.53	0.64	Significant
30	0.72	0.39	0.55	Significant
32	0.83	0.52	0.47	Significant
33	0.60	0.50	0.40	Significant
34	0.84	0.29	0.48	Significant
35	0.73	0.38	0.56	Significant
37	0.79	0.30	0.45	Significant
38	0.73	0.42	0.49	Significant
39	0.64	0.32	0.40	Significant
40	0.68	0.35	0.45	Significant

Referred to the third step test of Junior High School Medan city, implied that 36.4% of students were high ability, 49.3% were fair good and 15.3% were classified poor.

4. Conclusion

This experiment inferred that 33 items of questions that existed from the third step test is properly shifted to become a standard or benchmark to the mathematical characterics-based test.

Reference

[1] Ansari, B.L,(2009). *Concepts and Mathematics Application*. Banda Aceh: Foundation of Pen Publisher

[2] Asmin. (2004). *"Implementasi dari teori jawaban soal dalam ujian hasil belajar.* Medan Larispa, Indonesia.

[3] Azwar,Saifuddin. (2001). *Pengembangan Fungsi Ukur Prestasi Belajar*. Yogjakarta: Pustaka Pelajar Offset.

[4] Dahar, R.W. (1991). *Belajar Teory dan Pembelajaran.* Jakarta Erlangga.

[5] Djemari Mardapi. (2004). Penyusunan Tes Hasil Belajar. Yogyakarta: Program Pascasarjana Universitas Negeri YogyakartNaga,Dali S. (1992). *Teori Nilai Pengukuran*. Jakarta: Gunadarma University

[6] Groundlund,Norman E&Linn, Robert L. (1995). *Measurement and Assessment in Teaching*. New Jersey Prentice Hall, Inc.

[7] Hudoyo, H., (1990). *Matematika dan Pelaksanaannya di Depan Kelas*. Jakarta : DepDikbud.

[8] Hudoyo, H. (2005). *Pengembangan Kurikulum dan Pengajaran Matematika*. Malang Penerbit Universitas Negeri Malang.

[9] Nikko,Anthony J. (1989). Educational Assessment of Students. New Jersey. Merril, an imprint of Prentice Hall.

[10] Pophan,W. James (1995). *Educational Evaluation*. New York Mac Millan.

[11] Scufflebeam,D.I. (1972). *Educational Evaluation and Decision Making*, Peacock Publisher, Inc., Itasca, III.

[12] Suharsimi. (2004) *Arti Nilai Hasil Belajar*. Jakarta Bumi Aksara

[13] Wijaya, A. (2012). *Pendidikan Matematika Realistik Suatu Alternatif Pendidikan Pembelajaran Matematika*. Graha Ilmu: Yogjakarta.